RESEARCH ARTICLE

DESIGNING AN ENGLISH LANGUAGE PROFICIENCY TEST FOR FIRST-YEAR UNIVERSITY STUDENTS

HASSAN SAEED AWADH BA-UDHAN[1], WADDAH SALEH MOHAMMED MAHWARI[2]
[1]Seiyun University – Yemen, [2]Abyian University – Yemen
[1]Hassanbaudhan@yahoo.com; [2]wdhah1@hotmail.com

ABSTRACT

This study aims to design an English language proficiency test for measuring the first-year university students' proficiency in English. English proficiency test is designed to measure students' internal cognitive ability as well as their general communicative ability in English. The content of the English proficiency test is not based on the content or objectives of any courses which students taking the test may have followed. The test includes reading, vocabulary and grammar with multiple choice questions. A questionnaire is used to check the likeability of the test to the test-takers. The sample of the study consists of fifteen first-year students studying English as a foreign language at the College of Education, Seiyun University, Hadhramout, Yemen. Correlation, item facility, and item discrimination are used as analysis methods. Findings reveal that the reading section includes some very difficult and some very easy items, the vocabulary and grammar sections include many very difficult items and not even a single item which is very easy. Most students find the vocabulary items the most difficult to answer and the reason is found to be the distractors of these items are useless. Moreover, the study also finds out that the highest correlation is between reading and grammar sections, and the lowest correlation is between grammar and vocabulary and between vocabulary and reading. Some faulty items are improved by carrying modification either their stems or distractors. Finally, the study concludes that the test is useful and interesting for the target group as a proficiency test.

Key words: Test Design, English Proficiency Test, English Language Test Development.

Introduction

Many types of tests are used in English testing. According to Hughes, 1996 tests encompass various types based on the purpose of testing: placement test, diagnostic tests, achievement tests, and proficiency tests. Some tests are related to measuring the students' achievements in specific skill in English according to what they taught in the class, these types of tests are prepared by teacher and called achievement tests because they measure specific points/subjects in one skill. Other ones are called 'proficiency tests' which test everything connected with proficiency in English; such as TOEFL and IELTS used as a criterion for measuring students' proficiency in English. Some other tests aimed to place students at specific level of instruction, these are called placement tests. Tests with the purpose of discovering the weaknesses and strengths in students' level are called diagnostic tests.

In order to assess the quality of English, or the performance of the student in English, the teacher needs first to make clear why a target skill is evaluated. For example, in achievement tests, ignoring examining the aims of the course in advance leads to mismatch between what learners have taught in the course (the syllabus) and the content of the achievement test that are supposed to measure whether the learners have mastered the content of the course. That is to say, the lesson objective states that "at the end of the lesson, students will be able to describe orally the physical appearance of people", but as an achievement test the students are asked to describe several pictures of people in writing. The students here might successfully do the written description task, but this success is not indicative of their ability to orally describe people's physical appearance. This is a test of writing ability rather than a speaking ability, though it tests the same sub-skill (description of people).

1. Background to the study

The developed test in the current study is a proficiency test which aims to measure students' internal cognitive ability as well as their communicative ability. Harmer (2002) points out that a proficiency test gives a general picture of a student's knowledge and ability. Students' knowledge refers to grammar and vocabulary, while their abilities include listening ability, reading ability, writing ability, and speaking. Thus, the current study focuses on reading, vocabulary, and grammar. The three sections of the test (reading, vocabulary and grammar) measure students' internal ability to master the English language. These three areas are relevant to the students' academic needs that help them do well in their English classrooms. As stated by Abedi (2008) and Stephenson, Jorgensen, and Young (2004), students' academic English proficiency helps in their academic success in real classroom. Catering the test to the test-takers' needs assists test developers to predict how the test-takers use the language in future. It can be predicted, for instance, that these students definitely pass the leaving school exams (their need) based on the performance of Yemeni school students in a test that focuses on grammar and reading.

The current study investigates the relationship between vocabulary knowledge, grammar knowledge and the skill of reading in terms of the English-majoring EFL students. The relationship between language element (grammar and vocabulary) and language skills (listening, speaking, reading, and writing) is much under-researched. Few studies, as stated by Cheng and Matthews (2018), have examined the relationship between vocabulary knowledge and the reading and listening skills among L2 learners.

This study further gives language teachers and language testers insights on how to design valid and reliable tests. The solidness of language test is very crucial due to the decision taken based on test results. Giraldo (2018) concludes that language teachers make decisions based on assessment data and such decisions impact teaching and learning. This study provides a practical example for teachers to help them understand and improve their learner's language assessment practices. Sultana (2019) emphasizes that teachers' language need to understand how to apply their language assessment knowledge received in their training courses into their classroom practices.

1.1. Aims of the study

The aim of this paper is to design a proficiency test for assessing the first-year college students studying English as a foreign language.

1.2. Methodology

The current study is a mixed approach of quantitative and qualitative methods. The data collection instruments used in this study are proficiency test and questionnaire. A proficiency test is designed according to the stages of test development specified by Brown (2004) and Genesee & Upshur (1996). These stages include the test purpose, the test objectives and the test specification. Following the stages of test development ensures the validity and reliability of the test. Any test developer ignores even a single stage of test development; their test might not achieve its required purposes. In the current study, the proficiency test designed for the first-year college students takes into consideration all these stages of test development.

The language proficiency test for the target group aims to check students' ability to read and to be good at grammar and vocabulary of English. The test is a proficiency test with only fixed response items in three sections viz reading, vocabulary, and grammar (the items are contextualized). All items of the test are designed to be communicative because the main purpose from a language is to be used as a tool of communication. Thus, the test paper contains three sections: reading, vocabulary, and grammar.

The test items are also designed based on the test objectives stated below in order to ensure that the test items achieve their objectives and suit the level of the students. Besides, the test includes only multiple choice questions (MCQs) which are viewed to be the most challenging item type in terms of its designing and the most suitable for item analysis purposes.

1.3. Sample of the study

The sample of the study consists of fifteen first year students studying at the English Department at the College of Education, Seiyun University, Hadhramout, Yemen for the academic year 2018-2019. The students start to study English in a Yemeni context with English as a foreign language from seventh grade of primary education. All students come from a background where no exposure to English is. They don't use English in their social communication outside the college. They usually use English for academic purposes only.

1.4. Designing a proficiency test

Reading, vocabulary, and grammar are the focus of the designed test. The three sections of the test (reading, vocabulary and grammar) measure students' internal ability to master the English language. The students at this level are expected to have the competence in all the four basic skills of English and also both the elements of English viz grammar and vocabulary. In terms of reading they should be able to comprehend texts at both factual and inferential levels. They should be familiar with the words which are usually used in daily life contexts in academics. They should have the competence to use grammar correctly in written and spoken discourse. They should be able to write texts using the various devices of cohesion and coherence. They should be able to write well organized texts in particular formats, taking care of the appropriate content and vocabulary.

1.5. A proficiency test objectives

In general, assessed objectives must reflect what the test designer claims to have measured (Fives and DiDonato-Bames, 2013). In order to ensure this reflection, the following language objectives are written based on the target test-takers' assumed needs.

1.5.1. Reading objectives

- to test the different levels of comprehension of the test-takers in different text types viz narrative, and expository texts,

- to test the test-takers' ability to paraphrase information from the reading texts,

- to test the test-takers' ability to guess meaning of unfamiliar words in context,

1.5.2. Vocabulary objective

- to test the test-takers' ability to understand the use of vocabulary in context, and

1.5.3. Grammar objective

- To test the test-takers' ability to use tenses, the correct sentence structure, prepositions, possessive, countable and uncountable nouns, adverbs, phrasal verbs, and adjectives.

1.6. Test specification: Section-wise

Test specifications need to be thought of before starting creating the test itself. These specifications, as Brown (2004) mentioned, include 1) a broad outline of the test, 2) what skills you will test, and 3) what the items look like. Test specification table is created to ensure that each section and each item in each section measures exactly what it is supposed to measure content validity (see Appendix B for the table of test specification). Test

HASSAN SAEED AWADH BA-UDHAN & WADDAH SALEH MOHAMMED MAHWARI

specification table shows the test content and explain the procedure for assessing it. It also offers the relationship between the test objectives and the items or tasks of the designed test (Pitoniak, Young, Martiniello, King, Buteux, and Ginsburgh, 2009)

1.7. Application of test development principles on the designed test

The five principles for test development; namely, validity, reliability, authentic, backwash, and practicality; have been considered when designing the test under investigation. Each item in the test works out to be valid as all the items in the three sections are contextualized and based on interesting themes related to the students' background knowledge. Different subareas are covered in each of the three sections. For example: grammar subareas include tenses, the correct sentence structure, prepositions, possessive, countable and uncountable nouns, adverb, count and uncount, phrasal verb, etc.All these subareas are tested through two texts (contextualization). The three sections are all MCQs so the reliability of these items is very high in terms of scoring. All items of the test are contextualized so that items appear authentic and related to the target language domain of use to ensure that the test helps the test-takers to learn from the test and gain a positive backwash from it. The test paper has a time practicality; two hours, is enough for doing the test.

2.   Data analysis

Analysis of students' raw data indicates that the students have best performance in reading; whereas students' performance in grammar, compared to their performance in vocabulary, is better. However, such overall analysis is not enough to know in which section students performed better. The overall percentage of success of the whole test is 39 %. This finding indicates that the students in general have poorly performed in the test. The low performance in the test could not be attributed to the suitability of the test to the students' level in English. That is because, in the after-test questionnaire, many of the test-takers (60%) stated that the test was moderate in terms of easiness and difficulty. Few students (20%) said the test was easy, whereas other few students (20%) said that it was difficult. Moreover, the time allotted for the test could not also be the reason behind the students' low performance in the test for no single student stated that the test time was insufficient or less than they required; most students (69%) mentioned that the test time was appropriate. The rest of them (31%) declared that the time was more than needed. Furthermore, the instructions of the test, as most students (87%) confirmed in the questionnaire, were clear to them; the test instructions were not clear only for few students (13%). Thus, the major reason behind the students' low performance in the test should be the inclusion of faulty items or questions in the test. That is because about half students (47%), as the questionnaire data shows, declared that they disliked some questions and they attributed such dislike to their inability to understand the question itself. As far as the section-wise success is concerned, the students have the most achievement in the reading section (ms= 7.8), followed by the grammar section (ms= 3.53), and the least achievement in the vocabulary section (ms= 2.53).

2.1. Correlation between grammar, vocabulary, and reading sections

The correlation between the reading and grammar sections is 0.311 which is rather high and shows that students have performed rather consistently across the two sections, while the calculation between the grammar and vocabulary sections is rather low (0.104) which shows no consistent performance across the two sections. The correlation between the vocabulary and reading sections is (0.168) is rather low and shows that students have not performed very consistently across the two sections.

2.2. Item analysis

Item analysis includes three aspects: item facility analysis, item distractor analysis, and item discrimination analysis.

2.2.1. Item facility analysis

Item Facility analysis is useful in specifying which test item is too easy and which one is too difficult for the test-takers to do. The possible range of IF is 0.15 to 0.85 [higher value indicates too easy item; lower value indicates too difficult item]. However, here in this analysis the range that is accepted is 0.30 to 0.0.80; items that

are between 1 and 0.0.80 are too easy and items between 0.00 and 0.30 are too difficult. The IF indexes of the test conducted to the students are discussed below. The IF of each section is separately presented so that comparison among the three sections could be made.

a. Item facility in reading section

The following table shows the IF of each item in the reading section.

Table 1: Item facility in reading section

Subsections	I.	Items	1	2	3	4	5
		IF	0.33	1	0.40	0.73	0.47
	II.	Items	1	2	3	4	5
		IF	0.67	0.33	0.26	0.67	0.33

Table 1 shows that in the reading section, Item 3 in subsection II is very difficult for students to answer (0.26). The item as it appeared in the test is as follows:

- The experts and social workers suspect a big calamity at Deorari Lake. What made them think like this?

   a. They have found some cracks at the sides of the lake.

   b. The level of water in the lake is decreased.

   c. They have found some crack at the base of the lake.

   d. The lake belongs to an earthquake prone area.

Notice that the underlined option is the correct answer, yet most students fail to choose. The reason for the item difficulty could be the item itself is faulty. May be the distractor (s) are also possible correct answer(s). The item discrimination ability between the high and low performing groups could also be faulty.

However, item 5 in subsection I (IF=0.47) and item 2 (0.53) seems to be a perfect item for around half of students answered them correctly whereas the other half students answered them wrongly.

These findings show that in the reading section, item 3 in subsection II needs modification because it is very difficult. Similarly, item 2 in subsection I needs modification too because it is very easy and includes no much challenge to students.

b. Item facility in vocabulary section

The following table shows the IF of each item in the vocabulary section

Table 2: Item facility in vocabulary section

Subsections	III.	Items	1	2	3	4	5
		IF	0.00	0.13	0.47	0.27	0.27
	IV.	Items	1	2	3	4	5
		IF	0.20	0.53	0.07	0.33	0.27

Table 2 shows that in the vocabulary section most items are very difficult and that there is no single item very easy for students to answer. Item 1 in subsection III is very much difficult for none of the students got it correct. This finding shows that most items in the vocabulary section are very difficult justifies the previous findings that students' performance in the vocabulary section is the least among the performance of the three sections. Thus, students did not perform well in the vocabulary section because the items were very difficult.

HASSAN SAEED AWADH BA-UDHAN & WADDAH SALEH MOHAMMED MAHWARI

However, item 3 in subsection III (IF=0.47) and item 2 in subsection IV (IF=0.53) seem to be perfect items for around half of students answered them correctly whereas the other half students answered them wrongly. Item 3 in subsection IV as it appeared in the test is as follows:

- The word **monster** probably means:

  a. a very large and beautiful creature

  b. a very small and good-looking animal

  c. a kind of tree that bears fruit

  d. a very large and ugly creature

Item 2 in subsection V as it appeared in the test is as follows:

- The wise man took the youth for a _____ (2) through a garden.

  a. wander

  b. walk

  c. journey

  d. picnic

However, these two items may still have problems in terms of the distractors. The three distractors may not equally work as distractors. If one distractor is not at all chosen by students as correct answer, that distractor will be a bad distractor.

c. Item facility in grammar section

The table below shows the IF of each item in the grammar section

Table 3: Item facility in grammar section

Subsections	V.	Items	1	2	3	4	5
		IF	0.27	0.46	0.33	0.60	0.53
	VI.	Items	1	2	3	4	5
		IF	0.27	0.13	0.47	0.27	0.20

The table above shows that item 1 in subsection V is very difficult and that most items (1, 2, 4, and 5) in subsection VI in the grammar section are very difficult items. The table also shows that no item at all in the grammar section seems to be very easy for students to answer. Item 2 in subsection VI (0.13) is the most difficult in the section.

The finding that most items in subsection VI in the grammar section are very difficult to answer justifies the previous findings that students' performance in the grammar section is much less than their performance in reading section.

However, item 2 (IF=0.46) and item 5 (IF=0.53) in subsection V and item 3 in subsection VI (IF=0.47) seem to be perfect items for around half of students answered them correctly whereas the other half students answered them wrongly. However, these items may still have problems with their distractors.

2.2.2. Item Distractor analysis

Distractor analysis is a process aims to examine the results of the distractors to help identify good and poor distractors (Fulcher and Davidson(2007). To analyze distractors, a table prepared with the number of items as column headings and the different options as rows. Then the percentage of students who chose each option

is entered. The total percentage for all potions should be 1.The tables 10,11 and 12 below show which options are chosen by students in the three sections: reading, vocabulary and grammar. Notice that the shaded options are the correct answers and options with 0.00 value indicate that none of the students selected it as a correct option, whereas options with 1 value indicate that all students (100%) selected it as the correct option.

a.   Distractor analysis in reading section

The following table shows the distractor analysis of items in the reading section

Table 4: Distractor analysis in reading section

Subsections		Items \\ Options	1	2	3	4	5
	I.	A	0.60	0.00	0.20	0.07	0.47
		B	0.07	1	0.33	0.20	0.20
		C	0.00	0.00	0.07	0.73	0.27
		D	0.33	0 .00	0.40	0.00	0.06
	II.	Items \\ Options	1	2	3	4	5
		A	0.20	0.07	0.27	0.33	0.07
		B	0.00	0.33	0.27	0.67	0.20
		C	0.67	0.40	0.26	0.00	0.33
		D	0.13	0.20	0.20	0.00	0.40

The table above shows that in the reading section, item 3 in subsection II(see Appendix A for the test items) is a problematic item for it is very difficult for students to answer (IF=0.26). However, the item has no problems with its three distractors. Almost equal number of the students selected the three distractors as the correct answer, A (0.27), B (0.27) and D (0.20). This means that distractors do not need any modification.

On the other hand, table 4 above also shows that none of the students chose options A, C, and D in item 2 of subsection I (the three distractors) as the correct answer. This means the three distractors are useless. This item as previously discussed (see table 1) is very easy for it tests a low level of reading subareas which is finding out a local factual information in the text; such kind of question is usually easy for test-takers to answer; unless there's time limit for finding such kind of information in a text, the challenge is completely absent for scanning a text for specific information.

b.   Distractor analysis in vocabulary section

The table below includes distractor analysis of each item in the vocabulary section.

Table 5: Distractor analysis in vocabulary section

Subsections	III.	Items \\ Options	1	2	3	4	5
		A	0.67	0.40	0.07	0.27	0.13
		B	0.13	0.13	0.26	0.27	0.33

			0.00	0.07	0.13	0.13	0.27
		C	0.00	0.07	0.13	0.13	0.27
		D	0.20	0.40	0.47	0.13	0.20
	IV.	Items / Options	**1**	**2**	**3**	**4**	**5**
		A	0.20	0.20	0.27	0.13	0.27
		B	0.47	0.53	0.33	0.33	0.13
		C	0.13	0.27	0.33	0.33	0.27
		D	0.20	0.00	0.07	0.13	0.33

The table above shows that in the vocabulary section, item 1 in subsection III (see Appendix A for the test items) is a problematic item for it is very much difficult for students to answer (IF=0.00). The item has also a problem with distractor 'A'. Most students (0.67) selected distractor A as the correct answer. This means that distractor 'A' needs modification for it could be a possible correct answer, whereas the correct answer (option C) needs replacement or modification for it could be possibly wrong answer.

Item 2 in subsection IV has only one useless distractor (option D), whereas the other two distractors (options A and C) are useful for some students selected them as the correct answer. The item as indicated previously (see table 2) is a good item in terms of its IF (0.53); half students has got it correct and the other half have it wrong. Now it is clear that two distractors are useful (options A and C), whereas the third one (option D) is useless. Thus, a modification or a replacement of the useless distractor will help improve the item.

However, all other items in the vocabulary section, as shown in table 5, could be to some extent a good item. That is because the items have no useless distractors; at least one student has selected one of their three distractors as the correct answer. So far it is obvious that the three distractors of most of the items are to some extent useful. Thus, these items are relatively positive.

c. Distractor analysis in grammar section

The following table provides the distractor analysis of each item in the grammar section

Table 6: Distractor analysis in grammar section

Subsections		Items / Options	**1**	**2**	**3**	**4**	**5**
Subsections	**V.**	**A**	0.20	0.27	0.07	0.60	0.27
		B	0.13	0.46	0.07	0.00	0.53
		C	0.27	0.07	0.53	0.40	0.00
		D	0.40	0.20	0.33	0.00	0.20
	VI.	Items / Options	**1**	**2**	**3**	**4**	**5**
		A	0.27	0.07	0.00	0.40	0.46
		B	0.33	0.33	0.47	0.13	0.07
		C	0.40	0.47	0.13	0.27	0.20
		D	0.00	0.13	0.40	0.20	0.27

Item 4 in subsection V includes two completely useless distractors (options B and C) for none of students selected it as the correct answer. Therefore, these two distractors need modification or replacement. However, the item level of difficulty (IF=0.60) is moderate. This means there is no problem with the correct answer and only the distractors are faulty.

On the other hand, item 2 in subsection V (IF=0.46), as table 6 shows, could be to some extent a good item. That is because the item has no useless distractors; at least one student has selected one of its three distractors as the correct answer. The item as indicated previously (see table 3) is a good item in terms of its level of difficulty (IF=0.46).Based on these findings, the three distractors of the item appear, to some extent, useful. Thus, the item is relatively positive.

Item 5 in subsection V, as table 20 shows, has only one useless distractor (option C), whereas the other two distractors (options A and D) are useful; in a sense that some students selected them as the correct answer. The item as indicated previously (see table 3) is a good item in terms of its IF (0.53). Now it is clear that two distractors are useful (options A and D), whereas the third one (option C) is useless. Thus, a modification or a replacement of the useless distractor will make the item a much better item. The same analysis could be said for item 3 subsection VI for it almost has an ideal level of difficulty (IF=0.47) but one distractor is completely useless (option A). So such faulty distractor needs modification.

2.2.3. Item discrimination analysis

Apart from knowing how easy/difficult the items are, and how the distractors contributed to the item difficulty, it is important to know whether the items can discriminate between the low performing and the high-performing students.

a. Item discrimination in reading section

The following table provides item-wise responses of students in rank order of the highest, middle and lowest groups in reading section

Table 7: Item-wise responses in rank order of the highest, middle and lowest groups of reading section

Name	I. Items					II. Items				
	1	2	3	4	5	1	2	3	4	5
IF	0.33	1.00	0.40	0.73	0.47	0.67	0.33	0.27	0.67	0.33
IF UPPER	0.20	1.00	0.60	1.00	0.60	0.80	0.40	0.40	1.00	0.80
IF LOWER	0.20	1.00	0.40	0.40	0.20	0.40	0.20	0.20	0.40	0.20
ID	0.00	0.00	0.20	0.60	0.40	0.40	0.20	0.20	0.60	0.60

Table 7 above shows that in the reading section, items 1 and 2 in subsection I are not useful in discriminating performance of groups. The former item is a very difficult item (IF= 0.33); most students in the high group have got it incorrect (IF upper=20) and most students in the low group have also got it incorrect (IF lower = 0.20). The latter item, on the other hand, is a very easy item (IF=1.00); all students of both the high and the low groups got it correct (IF upper/lower=1.00).

b. Item discrimination in vocabulary section

The following table provides item-wise responses of students in rank order of the highest, middle and lowest groups in vocabulary section

HASSAN SAEED AWADH BA-UDHAN & WADDAH SALEH MOHAMMED MAHWARI

IJELR
www.ijelr.in

Table 8: Item-wise responses in rank order of the highest, middle and lowest groups of vocabulary section

Name	III.					IV.				
	Items					Items				
	1	2	3	4	5	1	2	3	4	5
IF	0.00	0.13	0.47	0.27	0.27	0.20	0.53	0.07	0.33	0.27
IF UPPER	0.00	0.20	1.00	0.40	0.00	0.60	1.00	0.20	0.60	0.20
IF LOWER	0.00	0.20	0.40	0.20	0.00	0.00	0.00	0.00	0.00	0.20
ID	0.00	0.00	0.60	0.20	0.00	0.60	1.00	0.20	0.60	0.00

Table 8 above shows that item 2 in subsection IV discriminated very well between students in the high group and those in the low group. The item has IF total of 0.53 which means it is an ideal item in terms of its difficulty level. Besides all students in the high group have got it correct (IF upper=1.00) and none of students in the low group have got it correct (IF lower = 0.00). Such item could be an ideal item and it could be said that the item suits the level of the target test-takers. The item (see table 5) has only one useless distractor (option D), whereas the other two distractors (options A and C) are useful for some students selected them as the correct answer. Thus, a modification or a replacement of the useless distractor will create a much better item.

Table 8 above also shows that most items, except item 2 in subsection IV, in the vocabulary section have not been able to discriminate and few of them discriminate moderately performance of groups because their IDs of all items are less than 0.80. Therefore, most items in the vocabulary sections are problematic and, as it is also proved in previous findings (see for example tables 1, 2, and 3), most students found the items in the vocabulary section are difficult to answer.

c.   Item discrimination in grammar section

The following table provides item-wise responses of students in rank order of the highest, middle and lowest groups in grammar section.

Table 9: Item-wise responses in rank order of the highest, middle and lowest groups of grammar section

Name	V.					VI.				
	Items					Items				
	1	2	3	4	5	1	2	3	4	5
IF	0.27	0.47	0.33	0.60	0.53	0.27	0.13	0.47	0.27	0.20
IF UPPER	0.20	1.00	0.40	0.80	1	0.00	0.00	0.40	0.40	0.60
IF LOWER	0.20	0.20	0.40	0.40	0.00	0.60	0.00	0.60	0.00	0.00
ID	0.00	0.80	0.00	0.40	1.00	-0.60	0.00	-0.20	0.40	0.60

Table 9 above shows that in the grammar section, items 1 and 3 in subsection VI and item 2 have not been able to discriminate performance of groups (IDs=0.00). These items are very difficult; the number of students in the high and low groups who answered each of these items is equal.

Item 5 in subsection VI has discriminated very well between students in the high group and those in the low group. The item has IF total of 0.53 which means it is an ideal item in terms of its difficulty level. Besides all students in the high group have got it correct (IF upper=1.00) and none of students in the low group have got it

correct (IF lower = 0.00). Such item could be an ideal item and it could be said that the item suits the level of the target test-takers. The item (see table 6 above) has only one useless distractor (option C), whereas the other two distractors (options A and D) are useful for some students selected them as the correct answer. Thus, a modification or a replacement of the useless distractor will make the item a much better item.

However, item 3 in subsection VII as previously discussed (see table 3) is an ideal item in terms of its difficulty level (IF=0.47) for around half of students answered them correctly whereas the other half students answered them wrongly. In addition, the item (see table 6 above) has only one bad distractor (option A), whereas the other two distractors (options C and D) are useful. But now after finding the ID of the item (-0.20), the item becomes much more problematic. Students from the low group (IF lower=0.60) who got the item correct has outnumbered students from the high group (IF upper=40) who got the item correct. This finding indicates that something went wrong with the item that makes the lower-performing group performs better than the higher-performing group. However, one possible reason could be the former group may have a chance of cheating or peeping to their seatmates during the test time.

Similarly, item 1 in subsection VII is completely faulty item (ID=-0.60). Most students from the low group (IF lower=0.60) got the item correct, whereas none of students from the high group (IF upper=0.00) got the item correct. This finding indicates that something went wrong with the item itself that makes the lower-performing group know the correct answer whereas the higher-performing group did not know it. However, another possible reason could be the former group may have a chance of cheating or peeping to their seatmates during the test time.

3.   Item improvement

In this section six of the problematic MCQs are improved based on the item analysis done in the previous part of the section. Six items are improved; two easy items and four difficult ones. The analysis above showed that the whole test includes only two easy items and these items existed in the reading section. Most MCQ items of the test are very difficult especially in the vocabulary and grammar sections. Thus, two difficult items from the vocabulary and another two difficult items from the grammar section are improved. First the reading section, then the vocabulary section and finally the grammar section.

3.1. Item improvement in reading section

In the reading section, the easies item is item 2 in subsection I. All students got it correct. The item as it appeared in the test is as follows:

- The first fifteen questions were easy for

   a. Ying-kee to do

   b. Wai-choy to do

   c. a few students to do

   d. all the students to do

One of the possible reasons that this item is very easy is that it includes no challenge at all. In a sense that the question or the item tests students' ability to find out local factual information in the text; such kind of question is usually easy for test-takers to answer. For the answer is located only in one position in the text. The correct answer (option b) as it appeared in the reading text in the test is as follows: "… Wai-choy picked up his pencil and began to write. He easily finished the first fifteen questions…". Students can easily recognize an option that is lifted up from the text as the correct answer.

Thus, the item is supposed to measure students' ability to answer inferential questions. The stem, therefore, needs to be modified for the stem is almost lifted up from the text. By changing the stem the previous correct answer (option b) becomes a distractor, whereas the previous first distractor (option a) becomes the correct answer. The change in the stem also makes the item requires inferential answer rather than the pre-modification factual-local answer. The modified item will be as follows:

- The first fifteen questions were difficult for

  a. Ying-kee to do

  b. Wai-choy to do

  c. a few students to do

  d. all the students to do

The item now will be difficult enough for the target group test-takers for the test proved that their overall level in English is poor. The modification in the stem now makes the three distractors useful though no change happened in the distractors, except that one distractor becomes the correct answer and the correct answer becomes a distractor.

3.2. Item improvement in vocabulary section

As it is proved in section II, the vocabulary section is the most difficult section in the test for students to do. It is also found in the previous part of this section that in the vocabulary section most items are very difficult and that there is no item at all that is very easy for students to answer (see table 8 and 5 above). Therefore, we will do some improvement to two difficult items in the vocabulary section.

The first item that needs improvement is Item 1 in subsection III. The item as it appeared in the test is as follows:

- Which of the following sets of words can be used together with **warm** (e.g. like 'warm day' in the passage)?

  a. weather, fire, fever

  b. water, mood, discussion

  c. weather, welcome, reception

  d. water, temper, friend

This item is very much difficult for none of the students got it correct (IF=0.00). The item has also a problem with distractor 'A' (see table 5). Most students (0.67) selected the distractor 'A' as the correct answer. This means that distractor 'A' needs modification for it could be a possible correct answer and the correct answer (option C) need replacement or modification for it could be possibly a wrong answer. The third distractor (option D) seems to be a possible correct answer.

The item is very difficult because it tests students' collocation knowledge. Thus, the kind of collocation included in the item might be above the level of the students; for this reason they all got the item wrong. Another reason is that the distractors (a, b, and d) might appear students as closer to the correct answer than the correct answer itself (option c). Thus, to make the item less difficult the three words in each option could be reduced into two words only. The modified item will be as follows:

- Which of the following sets of words can be used together with **warm** (e.g. like 'warm day' in the passage)?

  a. weather, fever

  b. water, mood.

  c. weather, reception

  d. water, fire

Notice that because the previous third distractor (option D) seems to be a possible correct answer, two words are reduced from it 'temper and friend' and the word 'fire' added to that distractor. The item now looks less difficult to the target group and it still includes some challenge that makes it difficult enough too.

HASSAN SAEED AWADH BA-UDHAN & WADDAH SALEH MOHAMMED MAHWARI

The second item in the vocabulary section that needs improvement is item 3 subsection IV (see table 5 above). The item is very difficult for most students got it incorrect (IF=0.07). The item as it appeared in the test is as follow:

- "…. Now pull out that one," said the old man _____ (3) to a bush…."

  [1] showing

  [2] looking

  [3] touching

  [4] pointing

Concerning the distractors of the above item, the three distractors (a, b, and c) have almost equally been selected by students as the correct answer; so they are very useful. The problem remains is the correct answer (option d) for the number of students who selected it as the correct answer is much less than the number of students who selected each of the three distractors as the correct answer (see table 5). One possible explanation of this finding is that students (the target group) are not aware of the prepositions that collocate with certain verbs such as, 'point to', and 'look at. In the stem the preposition 'to' collocates with the verb 'pointing' preventing the other distractors to be the correct answer. Another possible explanation is that the test-takers did not pay attention to the preposition 'to' in the stem which is actually a part of the text given to them.

Thus, if the first explanation is accepted as the reason for students' failure to pick up the correct answer, then the modification needs to be in the options as well as in the stem. The verb with preposition "point to" may be new for the students whose level in English, as their performance in test proved, is lower than expected. Thus, the correct answer needs to be a distractor and one of the distractors needs to work as a correct answer with some modification in the stem. The suggested modified item is as follows:

- "…. Now pull out that one," said the old man _____ (3) at a bush…."

  a. showing

  b. <u>looking</u>

  c. touching

  d. pointing

Now the item becomes less difficult for the previous correct answer (option d) becomes a distractor and the second distractor (option b) becomes the correct answer. The item still tests the subarea of prepositional verbs but the prepositional verb 'look at' will be more familiar to students (the target group) than the prepositional verb 'point to'. In the stem, the preposition 'to' is replaced with the preposition 'at'. This change in the stem and in the options makes the item less difficult and more cater to the level of target group.

3.3. Item improvement in grammar section

The grammar items, as it is proved in section 1.1 are more difficult for students to do than the reading items. Like the items in the vocabulary section, most items in the grammar section are very difficult that no item at all is very easy for students to answer (see table 3 above). Therefore, we will do some improvement to two difficult items in the grammar section.

Item 1 in subsection VI in the grammar section needs modification because it is a difficult item (IF=0.27). The three distractors of the item are all good distractors for the test-takers selected them as the correct answer (see table 6). So there is no problem with the distractors. The item (see table 9) has not discriminated well between students in the high and low groups (ID=0.00). This means the item is faulty. The item as it appeared in the test is as follows:

- **Sumanth:** I've lost my wallet.
  Inspector: Where did you _____ it?

   a. lost

   b. losing

   c. <u>lose</u>

   d. losed

The possible reason for the item being faulty is the subarea that the item tests which is main verb in an interrogative sentence in the past simple tense. Students may not know the form that a main verb takes in an interrogative sentence in the past simple tense. The students, as their performance proved in the test, have very low level in English. So the possible solution is to make the item little easier. Because the distractors have no problems at all, the change needs to happen in the stem. The modified item is as follows:

- **Sumanth:** I've _____ my wallet.
  Inspector: Where did you lose it?

   1. <u>lost</u>

   2. losing

   3. lose

   4. losed

The change that happens in the stem changes slightly the tested subarea. Instead of testing main verb form in an interrogative sentence in past simple tense, the item now tests the main verb form in a declarative sentence in present perfect tense. The structure of declarative sentences is usually easier for students to understand than the structure of interrogative sentences. The correct answer in the previous item becomes a distractor in the modified items (option c), whereas the first distractor in the previous item becomes the correct answer in the modified item (option a) Thus, the item now will be better and less difficult for the test-takers.

The second item that would be modified is item 2 in subsection VII. The item is very difficult for most students to answer (IF=0.13). The item as it appeared in the test is as follows:

- … Someone with bad breath should drink_____ of water because dry mouth can also be a cause….

   a. very much

   b. more

   c. some

   d. <u>plenty</u>

The distractors of the item are all useful ones. So there is no problem with them as distractors. The reason of the difficulty of the item could be the word 'plenty' (the correct answer); it might be new word for the students whose level in English, as their overall performance (39%) in the test proved, is generally weak. The word 'plenty' could work well if it is replaced with the phrase 'a lot'. Thus, the modified item is as follows:

- … Someone with bad breath should drink _____ of water because dry mouth can also be a cause….

   a. very much

   b. more

   c. some

   d. a lot

By replacing the word 'plenty' with the phrase 'a lot', the item still tests the same subarea of grammar (a quantifier + of) and the item becomes less difficult. The quantifier 'a lot of' will be more familiar to the target

HASSAN SAEED AWADH BA-UDHAN & WADDAH SALEH MOHAMMED MAHWARI

group than the quantifier 'plenty of'. Such kind of change in the correct answer will make the item better and more suitable to the test-takers' level in English.

Conclusion

The target group finds the test interesting and useful. The multiple choice items have to be new experience for the test-takers who are used to only answer open-ended questions. The test-takers learn to improve their English from the test. For instance, they learn how reading texts of their interest help them understand better. All items are contextualized and communicative in order to draw the test-takers' attention to the importance of learning the language skills and elements in contexts and not in a discrete manner. Communicative and contextualized test items must be another new experience from which they learn to improve their language. Thus, the test is a source for learning more than for evaluating the test-takers' ability in English. To sum up the findings of the IF of the items in the three sections prove that all the three sections have faulty items. Whereas the reading section includes some very difficult and some very easy items, the vocabulary and grammar sections include many very difficult items and not even a single item which is very easy. This finding justifies the previous finding that students' least success is in the vocabulary and grammar sections compared to their success in reading section.

References

Abedi J. "Measuring Students' Level of English Proficiency: Educational Significance and Assessment Requirements." Educational Assessment, 13(1); (2008): pp 193-214.

Bachman L. F., and Palmer A. S. Language Testing in Practice. Oxford: Oxford University Press, 1996).

Brown D. H. "Designing Classroom Language Tests." Language Assessment: Principles and Classroom Practices. Ed. Brown D. H. New York: Longman, 2004. 42-65.

Cheng J., and Matthews J. "The Relationship Between Three Measures of L2 Vocabulary Knowledge and L2 Listening and Reading." Language Testing, *35*(1); (2018): pp 3-25.

Fives H. and DiDonato-Bames N. Classroom Test Construction: The Power of a Table of Specifications. Practical Assessment, Research & Evaluation, 18(3); (2013): pp 1-7.

Fulcher G. and Davidson, F. Language Testing and Assessment: An Advanced Resource Book. London: Routledge, 2007.

Genesee F., & Upshur J. A. "Objectives-referenced testing." Classroom Evaluation in Second Language Education. Genesee F. and J. A. Upshur. Cambridge: Cambridge Press University, 1996. 158-175.

Giraldo F. "Language Assessment Literacy: Implications for Language Teachers." *Profile: Issues in Teachers´ Professional Development,*20(1); (2018): pp 179-195.

Hanner J. The Practice of English Language Teaching. Malaysia: Pearson Educated Limited, 2002.

Hughes, A. Testing for Language Teachers. Cambridge: Cambridge University Press, 1996.

Pitoniak M. j., Young J. W., Martiniello M., King T. C., Buteux A., and Ginsburgh M. Guidelines for the Assessment of English Language Learners. Princeton, NJ: Educational Testing Service, 2009.

Stephenson A., Jorgensen M. A., and Young M. J. Assessing English Language Proficiency: Using Valid Results to Optimize Instruction. Paper presented at the 81st Annual Meeting of the California Educational Research Association (CERA, November 14, 2003, Revision 2, January 2004).

Sultana N. Language Assessment Literacy: An Uncharted Area for the English Language Teachers in Bangladesh. Language Testing in Asia, 9(1); (2019): pp 2-14.

Appendix A

Test paper and answer script samples

Date: Class:	PROFICIENCY TEST IN ENGLISH LANGUAGE	Maximum marks: 50 Allotted time: 2 hours

Name: _____   Roll No. _____

- Read all the instructions carefully before you begin to answer the questions.
- Write your name and roll number in the space provided above.
- Please check that this question paper contains 6 printed pages.
- Please check that this question paper contains 9 questions.

Instructions:

    i.   *This paper consists of three sections:*

        Section A   Reading   20 marks

        Section B   Vocabulary   15 marks

        Section C   Grammar   15 marks

    ii.   *Attempt **all** questions.*

    iii.   *You have to choose the correct answer (A, B, or C) and encircle it.*

Section A: Reading

I.   **Read the following passage.**   **10 Marks**

    The Abraham Prize

'Next Tuesday, we are going to have a special examination in English,' said the teacher. 'Many years ago, there was a boy at this school called William Abraham. That boy passed out of school and became doctor. Now, every year, Dr Abraham gives a prize to the student who gets the highest marks in a special English exam.'

On Tuesday morning everyone was ready for exam. 'You have to answer twenty questions in one hour,' the teacher said. 'And remember, you must not cheat.' The students began to do the exam. Wai-choy picked up his pencil and began to write. He easily finished the first fifteen questions.

Just then, piece of folded paper landed on his desk. Wai-choy looked round. He saw one of his classmates, Ying-kee, biting his pen and looking at him. Wai-choy opened the note. It read: *What is the answer to Question 11?*

Wai-choy looked angrily at Ying-kee and shook his head. He folded the note and threw it back. 'What are you doing, Wai-choy?' the teacher asked. He walked up and picked up the note from Ying-kee's desk. He read it and said, 'I told you not to cheat. Give me your paper and leave the room, Wai-choy. I can't let you finish the exam.'

Wai-choy walked out with tears in his eyes. He got on his bicycle and rode very slowly along the road with his head down. And before he knew it he had crashed into an old man. Both of them fell down.

'I'm very sorry, sir,' Wai-choy said. 'I wasn't looking.'

'That's alright,' said the old man

'let me help you get up,' said Wai-choy.

'Thanks. But tell me, why are you not in school?' asked the old man.

HASSAN SAEED AWADH BA-UDHAN & WADDAH SALEH MOHAMMED MAHWARI

'I was sent out of school, sir.' Wai-choy said. Soon he was telling the old man all about Ying-kee's note. The old man listened patiently. 'Can you prove to your teacher that you didn't write the note?' he asked.

Just then Wai-choy noticed a large pen in the old man's pocket. 'Oh yes, I can prove that I didn't write that note! He cried. Wai-choy rode like the wind back to school. When he entered the classroom, his teacher was marking the exam papers.

Sir, I didn't cheat in the exam,' said Wai-choy. 'The note was written with a pen and I always use a pencil. In fact, I don't have a pen. If you look at the handwriting in the note, sir, you'll see that it's not my handwriting. In any case, I had already answered.

The teacher checked the Wai-choy's answer paper and the note. He said, 'I'm very sorry, Wai-choy. I made mistake. Please forgive me. Now sit down and complete the exam.'

Finally Wai-choy got the highest marks in the exam. On School Day, when Wai-choy walked up to receive his prize from Dr Abraham, he was pleasantly surprised. Dr Abraham was the old man he had knocked down with his bicycle.

Now, based on your understanding of the passage, answer the following questions.

1. Who gives the prize to the student who gets the highest marks in a special English exam?

    a. The English teacher of the school

    b. The principal of the school

    c. The secretary of the school

    d. The former student of the school

2. The first fifteen questions were easy for

    a. Ying-kee to do

    b. Wai-choy to do

    c. a few students to do

    d. all the students to do

3. Why do you think Wai-choy was surprised while receiving the prize?

    a. He did not expect that he would win the prize.

    b. He was receiving the prize from Dr Abraham.

    c. Dr Abraham was a famous person whom everyone knows.

    d. Dr. Abraham was the man who Wai-choy hit with bicycle.

4. Which one of the following statements is true according to the passage?

    a. Ying-kee's note was written with a pencil

    b. The handwriting in the note was Wai-choy's

    c. Wai-choy would always use the pencil

    d. All the students were supposed to use the pencil

5. The teacher did not allow Wai-choy to take the exam because

    a. he saw Wai-choy throwing the note

    b. Wai-choy was not well

    c. he wanted Wai-choy to meet Dr Abraham

HASSAN SAEED AWADH BA-UDHAN & WADDAH SALEH MOHAMMED MAHWARI

    d. he did not want Wai-choy to win the prize

II. **Read the following passage**   **10 Marks**

    Land Slides

Experts and social workers at Ukhimath opine that the present catastrophe which has devastated 21 villages in this area was not a natural one but a man-made disaster. They feel that these deaths could have been averted if several callous decisions had not been taken by both the villagers and the district administration. They fear that a bigger calamity can occur at one of the biggest lakes in the area, Deorari Lake, since cracks have already developed at several places below the lake.

Mr Hanumant Rawat, a social worker who is also helping in relief operations, blames the large-scale felling of trees for the recurrent landslides in the region. "Trees help to retain the water and therefore prevent the sudden rush of water and soil as has happened on this occasion. But both the villagers and the government have consistently encouraged felling of trees which has resulted in total erosion of the top soil. Since it rains heavily in the Alpine layers and there are no trees to prevent the erosion, the massive landslides are the direct result of reckless felling of trees," he added. He warned that the entire Rudraprayag area would be submerged in floods if a similar calamity occurred in Deorari Lake.

Another reason for the sudden increase in landslides was terrace farming. "Villagers, out of greed are unscientifically changing the pattern of cultivation and cutting trees for terrace farming. This requires vast tracts of land to be denuded of forests and also requires a lot of water. This is the reason why landslides often occur in areas which have large "terrace farms," said Mr Dharmesh Kumar, a social worker who has done research in changing cultivation patterns in the region and their impact on the environment. This is the reason why there are a large number of rivulets in the Ukhimath region during rainy season. "Since there are no trees water starts flowing from anywhere. These small rivulets were not seen earlier in such a large number since the trees would prevent the flowing down of water," he added.

Now, based on your understanding of the passage, answer the following questions.

  1. Choose the odd word

        a. Disaster

        b. Calamity

        c. Cultivation

        d. Catastrophe

  2. The two main reasons for the repeated landslides are

        a. terrace farming and cracks below lakes

        b. terrace farming and felling of tress

        c. felling of trees and cracks below lakes

        d. cracks below lakes and villagers' decisions

  3. The experts and social workers suspect a big calamity at Deorari Lake. What made them think like this?

        e. They have found some cracks at the sides of the lake.

        f. The level of water in the lake is decreased.

        g. They have found some crack at the base of the lake.

        h. The lake belongs to an earth quake prone area.

  4. Who among the following (according to the passage) considers the changing cultivation pattern as one of the reasons of landslides?

    a. Mr Hanumant Rawat

    b. Mr Dharmesh Kumar

    c. The government

    d. The villagers

5. The author calls the landslide as a "man-made disaster" because

    a. the landslide killed many people of the area.

    b. the landslide was not a natural one but created under laboratory conditions

    c. the landslide was a result of the decisions taken by the people

    d. the experts (who are humans) did research on the landslide

Section B: Vocabulary   **10 Marks**

III. **In the following passage there are five underlined words. At the end of the passage there are questions based on each of these words. Read the passage and answer the questions that follow it.**

5 x1= 5 Marks

Near the close of an exceedingly **warm** day, I was sitting, book in hand, at an open window, commanding a view of a distant hill that had been denuded of its trees by a land-slide. I looked up from the book and **gazed** at the naked face of the hill and at an object--- at some hideous living **monster** that rapidly made its way from the **summit** to the bottom, disappearing in the dense forest below. When I first saw this creature, I doubted my **sanity**---or at least the evidence of my own eyes; and many minutes that I was neither mad nor dreaming.

1. Which of the following sets of words can be used together with **warm** (e.g. like 'warm day' in the passage)?

    a. weather, fire, fever

    b. water, mood, discussion

    c. weather, welcome, reception

    d. water, temper, friend

2. Find the odd word out from the following :

    a. gaze

    b. blink

    c. peep

    d. look

3. The word **monster** probably means:

    e. a very large and beautiful creature

    f. a very small and good-looking animal

    g. a kind of tree that bears fruit

    h. a very large and ugly creature

4. Which of the following can be a synonym of the word '**summit**'?

    a. peak

   b.   base

   c.   edge

   d.   centre

 5.   Choose the opposite of the word '**sanity**.'

   a.   dissanity

   b.   unsanity

   c.   insanity

   d.   illness

IV.   **Fill in the blanks in the passage given below. Choose the words from the options given for each blank at the end of the passage.   5 x1= 5 Marks**

A wealthy man requested an old wise man to _____ (1) his son away from his bad habits. The wise man took the youth for a _____ (2) through a garden. Stopping suddenly he asked the boy to pull out a tiny plant growing there.

The youth held the plant between his thumb and forefinger and pulled it out. The old man then asked him to pull out a slightly bigger plant. The youth pulled hard and the plant came out, roots and all. "Now pull out that one," said the old man _____ (3) to a bush. The boy had to use all his strength to pull it out.

"Now take this one out," said the old man, indicating a guava tree. The youth _____ (4) the trunk and tried to pull it out. But it would not budge. "It's impossible," said the boy, panting with the effort.

"So it is with bad habits," said the sage. "When they are young it is easy to pull them out but when they _____ (5) hold they cannot be uprooted."

The session with the old man changed the boy's life.

1.	**a.**	wean	**b.**	teach	**c.**	train	**d.** help
2.	**e.**	wander	**f.**	walk	**g.**	journey	**h.** picnic
3.	[5]	showing	[6]	looking	[7]	touching	[8] pointing
4.	**a.**	caught	**b.**	seized	**c.**	grasped	**d.** grabbed
5.	**a.**	take	**b.**	make	**c.**	have	**d.** get

                    Section C: Grammar   10 Marks

V.   **Fill in the blanks in the passage given below. Choose the words from the options given for each blank at the end of the passage.   5   x1   =   5 Marks**

Sumanth: Good morning, inspector.
Inspector: Good morning, what can I do for you?
Sumanth: I've lost my wallet.
Inspector: Where did you _____ (1) it?
Sumanth: In the bus, I think. My pocket was probably picked.
Inspector: Well, can you describe it?
Sumanth: It is brown _____ (2). It is new, I bought it just yesterday.
Inspector: Is there any money in it?
Sumanth: Oh, yes. Five-hundred rupees _____ (3) my fees.

HASSAN SAEED AWADH BA-UDHAN & WADDAH SALEH MOHAMMED MAHWARI

Inspector: What else did you have in it?

Sumanth: Well, my address, my _____ (4) photograph and _____ (5) stamps.

Inspector: Is this your wallet?

Sumanth: Good heavens, yes! Where did you find it?

Inspector: I did not find it. Someone else did. Here is his name and address. Write and thank him.

Sumanth: I will. And thank you, sir.

1.	**e.**	lost	**f.**	losing	**g.**	lose	**h.**	losed
2.	**a.**	little leather with a pocket	**b.**	with a little leather pocket	**c.**	a pocket with little leather	**d.**	leather with a little pocket
3.	**a.**	of	**b.**	to	**c.**	from	**d.**	for
4.	**a.**	mother's	**b.**	mothers	**c.**	mother	**d.**	mothers'
5.	**a.**	little	**b.**	some	**c.**	much	**d.**	any

VI.   Fill in the blanks in the passage given below. Choose the words from the options given for each blank at the end of the passage.   5 x1 = 5 Marks

The problem of bad breath is_____ (1) caused when bacteria in the mouth break down food particles into sulphurous or other volatile compounds. Someone with bad breath should drink_____(2) of water because dry mouth can also be a cause.

Older people and snorers are particularly at risk of_____(3) from a dry mouth. Brushing the_____(4) helps but not enough to prevent bad breath. If proper oral hygiene and drinking lot of water do not get rid of bad breath, the _____ (5) person should have a medical examination.

1.	**a.**	usually	**b.**	casually	**c.**	naturally	**d.**	orally
2.	**a.**	very much	**b.**	more	**c.**	some	**d.**	plenty
3.	**a.**	having	**b.**	suffering	**c.**	aching	**d.**	brushing
4.	**a.**	teeths	**b.**	tooth	**c.**	teeth	**d.**	toothpaste
5.	**a.**	affective	**b.**	affecting	**c.**	affected	**d.**	affect

Appendix B

Table : Section-wise test specification

Section A : Reading							
Answer Key: (See Appendix 1)							
Subsection	**Text type**	**Theme**	**Item type**	**Length of input**	**Length of response**	**No of items & score**	**Subarea tested**
1.	Story	Honesty Reward	fixed response	500 words + MCQs	------	5 items 1x5=5	1. local factual 2. factual-assimilation 3. global inferential 4. factual 5. inferential

HASSAN SAEED AWADH BA-UDHAN & WADDAH SALEH MOHAMMED MAHWARI

2.	Expository/ informative	environmental awareness	fixed response	300 words + MCQs	-------	5 items 1x5=5	1. meaning 2. factual assimilation 3. inferential 4. factual assimilation 5. inferential

Section B : Vocabulary

Answer Key: (See Appendix 2)

Sub-section	Theme	Length of input	Length of response	Item type	No. of items & Score	Subarea tested
3.	Fear/surprise (paragraph)	70 words + MCQs	-	fixed response	5 items, 5x1= 5 Marks	6. collocation, 7. meaning 8. meaning 9. synonym 10. antonym
4.	Discipline	150 Words + MCQs	-	fixed response	5 items, 5x1= 5 Marks	1. words in context 2. words in context 3. words in context 4. words in context 5. words in context

Section C : Grammar

Answer Key: (See Appendix 3)

Subsection	Text type	Theme	Length of input	Length of response	Item type	No. of items & Score	Subarea tested
6.	Dialogue	Pocket-picking	80 Words + MCQ items	-	fixed response	5 items, 5x1= 5 marks	1. tenses 2. the correct sentence structure 3. prepositions 4. possessive 5. countable and uncountable nouns
7.	Newspaper article	Health	80 words + MC items	-	- fixed response	5 items, 5x1= 5 marks	1. adverb 2. countable and uncountable 3. phrasal verb 4. noun number 5. adjective

Appendix C

Feedback questionnaire sample (unfilled + filled)

1.   Did you like the test?   (Yes/No)
2.   If yes/no, why?
3.   Were the instructions clear?   (Yes/ No)
4.   How easy /difficult was the test?
     a.   Easy
     b.   Moderate
     c.   Difficult
     d.   Very difficult
5.   Which of the three sections was easiest?
     a.   Reading
     b.   Grammar
     c.   Vocabulary
6.   Which of the three sections was most difficult?
     a.   Reading
     b.   Grammar
     c.   Vocabulary
7.   Have you ever come across such kind of questions (questions with four options) before? Yes/No
8.   Did you dislike any of the questions? If yes, why?
9.   Did any of the questions surprise you? If yes, why?
10. The time allotted for writing the test was—
               a.   Less than needed
               b.   As per the need (Appropriate)
               c.   More than needed

HASSAN SAEED AWADH BA-UDHAN & WADDAH SALEH MOHAMMED MAHWARI